How Do You Convert Your Legacy Content to DITA?

Online Survey Results

Table of Contents

Overview 3

Survey Results 4

Legacy Content Format (Question 1) 4

The State of the Legacy Documentation (Question 2) 6

Amount of Legacy Content (Question 3) 7

Types of Legacy Content (Question 4) 8

Restructuring and Rewriting Legacy Content (Question 5) 9

DITA Information Types (Question 6) 11

Challenges (Question 7) 12

Who Was in Charge (Question 8) 14

Conversion Tools (Question 9) 16

Overall Impression (Question 10) 18

What If... (Question 11) 20

Final Comments (Question 12) 23

About DITATOO 25

(c) Copyright 2013 DITATOO. All rights reserved.

Overview

One of the must-do steps in many DITA implementation processes is to convert the existing content to DITA.

From October 2012 through February 2013, we, at DITAToo (http://www.ditatoo.com), were conducting an online survey in order to collect information on experience that documentation teams gained when converting their legacy content to DITA.

To conduct the survey and analyze the results, we used SurveyMonkey (www.surveymonkey.net).

The link to the survey was published in technical writing mailing lists, such as dita-users and HATT, and various LinkedIn groups. We have received 86 responses.

The survey included twelve questions:

1. In what format is your legacy content stored?
2. If your legacy content is stored in an unstructured format, such as MS Word or unstructured Framemaker, which statement best describes your situation?
3. What is the amount of your legacy content to be converted (in pages)?
4. What kind of information does your legacy content contain?
5. Do you need to rewrite or restructure your legacy content to make it fit DITA in terms of content organization and writing principles? Please select the statement that best describes your situation.
6. To which DITA information types do you need to convert your legacy content?
7. What challenges did you face (or are facing now) when converting your legacy content to DITA?
8. By whom the conversion was done (or is being performed)?
9. What tool(s) do you use to convert your legacy content to DITA?
10. What is your overall impression from the conversion?
11. If you would do the conversion now, when you already know all pitfalls and issues, what would you do differently?
12. Please share with us any comments or thoughts that you have on the conversion

Respondents could post additional comments to each question.

(c) Copyright 2013 DITAToo. All rights reserved.

Survey Results

The following sections provide detailed information on the results that we have received.

Legacy Content Format (Question 1)

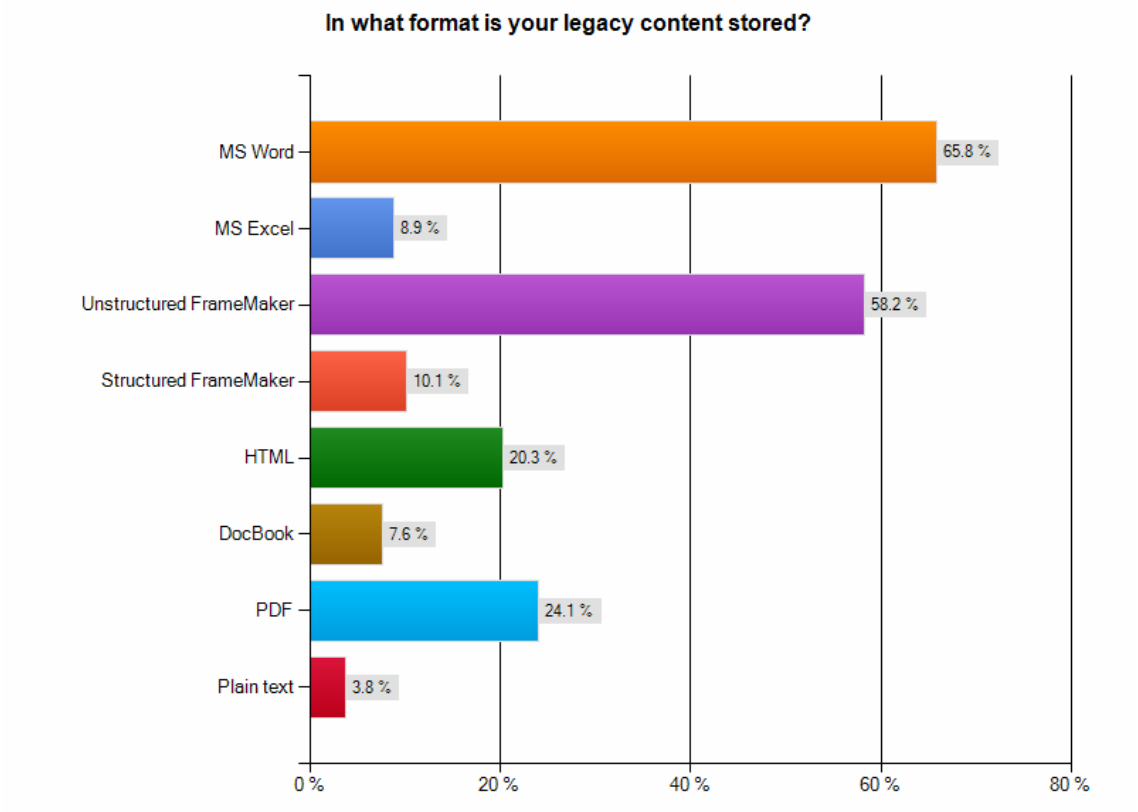
First of all, we asked our respondents in what format their legacy content is stored (multiple answers could be selected).

Although in the first days after the survey was opened, unstructured FrameMaker was leading the race, the final score shows that the majority of respondents have their legacy documentation in MS Word (**65.8%**). Unstructured FrameMaker is used by **58.2%**.

PDF was indicated as a legacy content format by **24.1%**. In some cases, this means that the original source is lost, and documentation teams have to work with PDFs.

20.3% keep their legacy documentation in HTML, which might mean they are using tools, like Madcap Flare or Adobe Robohelp. Some respondents noticed in the comments that they have legacy content in InDesign.

This chart shows a graphical representation of the answers:



(c) Copyright 2013 DITATOO. All rights reserved.

Optional Comments

"RoboHelp version 9 MS Word"

"Robohelp"

"By HTML I mean RoboHelp projects; mostly Web Help with some JavaHelp"

"Adobe InDesign and some really old Quark files"

"Madcap Flare"

"InDesign"

"All our documents are published as PDF produced from Word. We have most of the orininal Word documents."

"Primarily authored in InDesign, published as PDFs"

"Quark"

"3B2, XyASCII, Miles"

"Other XML formats IBM DITA - DITA before standardisation"

"Proprietary XML"

"iMap XML"

"MediaWiki SharePoint"

"WIKI"

"PowerPonit (training information)"

"Acrobat Epic Editor"

"Various graphics formats Specialized databases"

"Help & Manual"

(c) Copyright 2013 DITATOO. All rights reserved.

The State of the Legacy Documentation (Question 2)

Almost any conversion from unstructured formats is based on styles. As long as legacy documents are styled consistently and no manual formatting is applied, there is a good chance that the conversion will be fairly smooth.

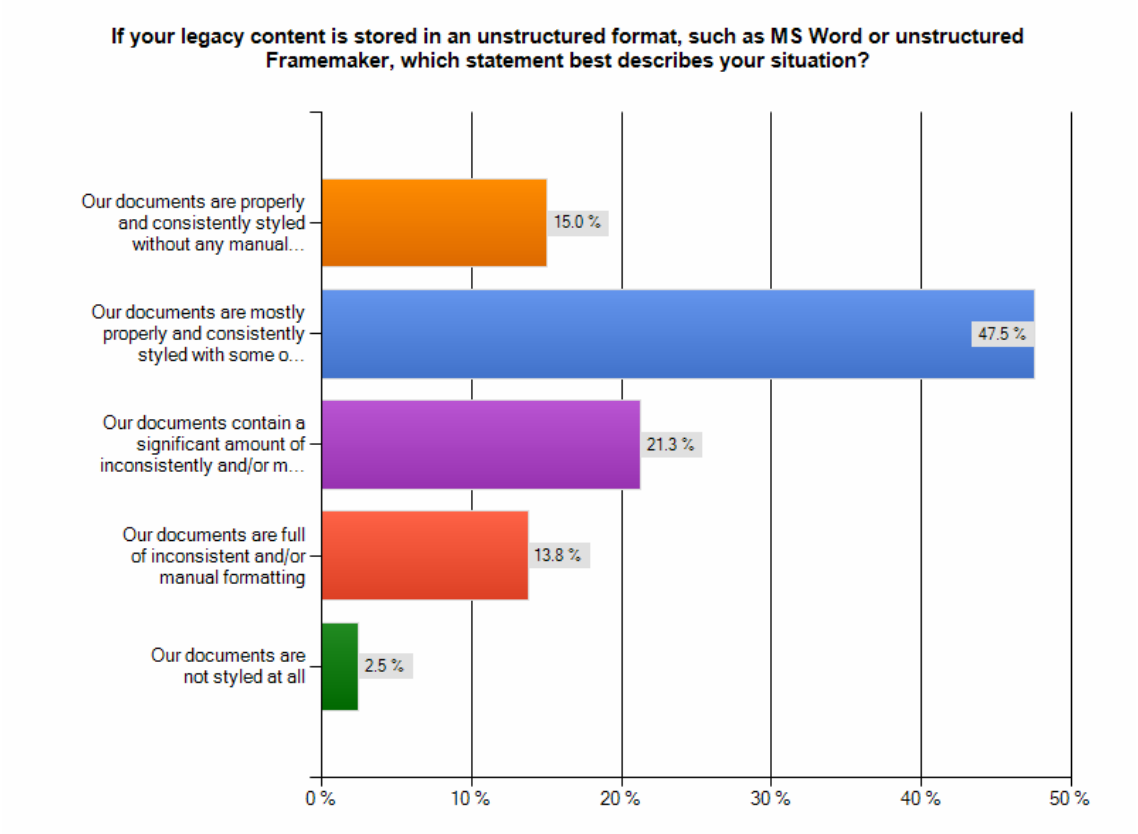
To understand if the legacy content is consistently formatted, we asked our respondents to choose the statement that best describes their situation.

The vast majority (**47.5%**) answered that their documents are mostly properly and consistently styled with some occasional manual formatting. **21.3%** stated that their documents contain a significant amount of inconsistently and/or manually formatted content.

The documents are properly formatted without any manual formatting for **15%**. This goes quite close with the answer coming from another edge: **13.8%** have legacy documents full of manual and inconsistent formatting. Finally, **2.5%** said that their legacy documents are not styled at all.

After drilling down the individual responses, we did not find any clear correlation between the amount of inconsistent formatting and authoring tools used to create this content. Users of unstructured FrameMaker tend to have a little bit more consistent formatting, but due to a limited number of responses, we cannot identify this as a clear trend.

This chart shows a graphical representation of the answers:


If your legacy content is stored in an unstructured format, such as MS Word or unstructured Framemaker, which statement best describes your situation?
Our documents are properly and consistently styled without any manual...
15.0 %
Our documents are mostly properly and consistently styled with some o...
47.5 %
Our documents contain a significant amount of inconsistently and/or m...
21.3 %
Our documents are full of inconsistent and/or manual formatting
13.8 %
Our documents are not styled at all
2.5 %
0 %
10 %
20 %
30 %
40 %
50 %

(c) Copyright 2013 DITATOO. All rights reserved.

Optional Comments

"Conversion was done in about 4 months, but it was a debacle because of the way it was handled by upper management."

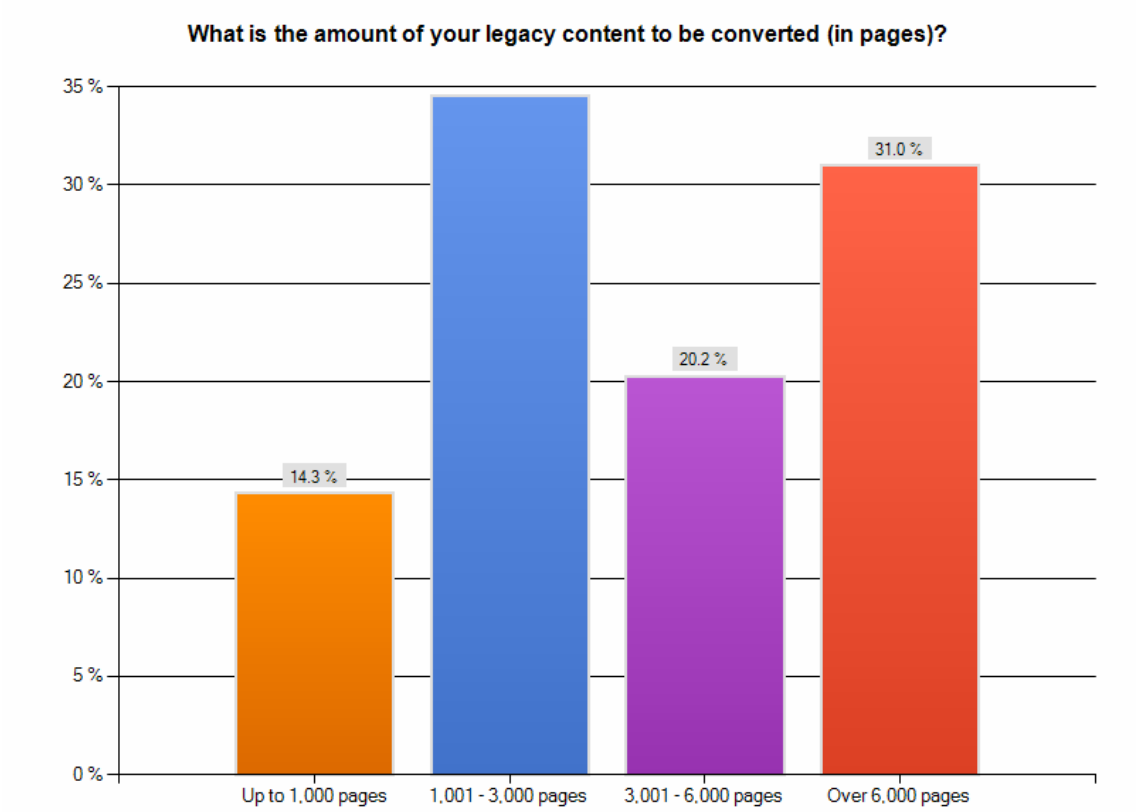
"We have multiple templates. Some promote consistency of content. Another does noy."

"Our styles are fine. It's the "non-topicised" nature that is the real issue."

Amount of Legacy Content (Question 3)

Most of the respondents estimated the amount of their legacy content in the range between 1,001 and 3,000 pages (**34.5%**). This number is followed by those writers whose content is over 6,000 pages (**31%**).

This chart shows a graphical representation of the answers:



We have to keep in mind, however, that the key factor in estimating the amount of content to be converted is whether you calculate the entire content accumulated over the years or just the content you actually need to have in DITA.

For example, we had a customer whose DITA implementation plan involved conversion of the legacy content. The initial page count was over 8,000 pages. However, after we investigated

(c) Copyright 2013 DITATOO. All rights reserved.

their content, we found that a lot of content was outdated and will never be used again. In addition, we also realized that there was a significant amount of duplicative content that should be rewritten before the conversion. In the end, we had to convert just about 4,000 pages, that is twice as less as the original estimate.

Optional Comments

"Varies from under 1,000 to over 6,000 "

"Not everything needs to be converted to DITA - estimating reusable portion only"

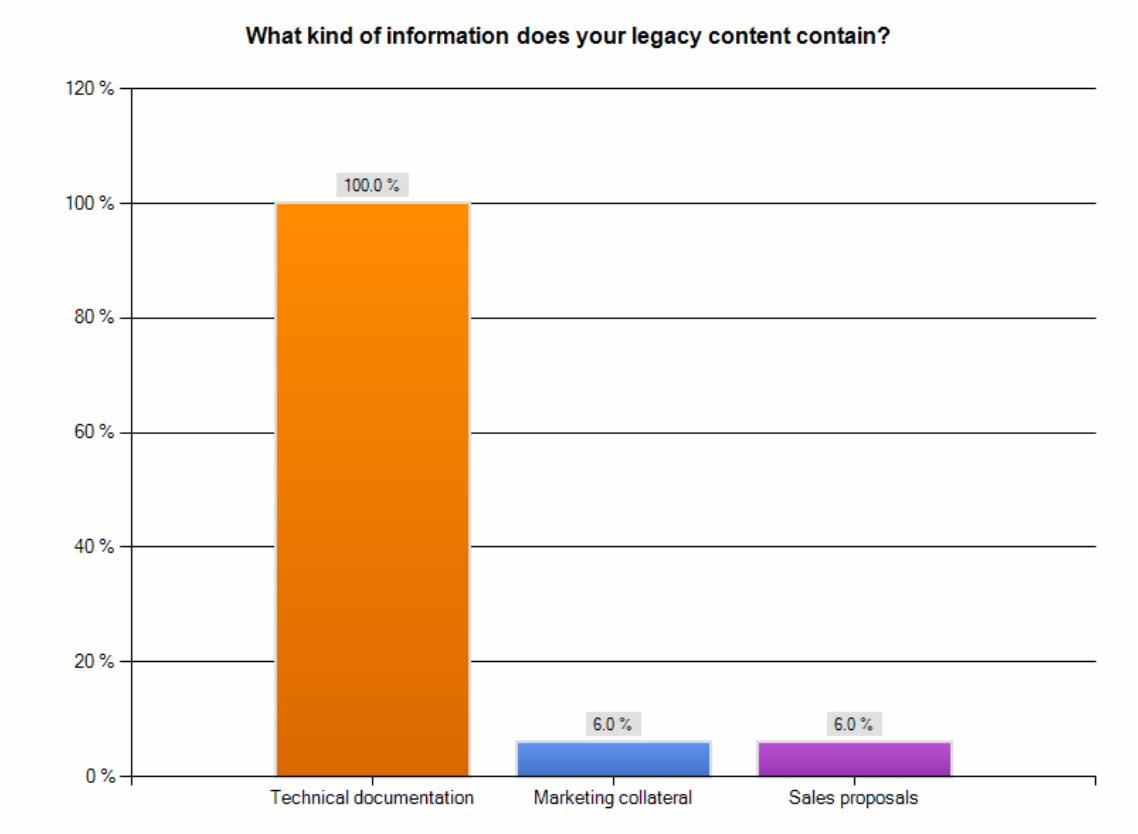
"Probably 80% of the individual projects in the 100-300 page range, probably 20% in the 3000 to 8000 page range."

"Really frighteningly huge."

Types of Legacy Content (Question 4)

We asked about the types of the information that the legacy content contains (multiple answers could be selected). Not surprisingly that **100%** of respondents said it is technical documentation with a slight portion of marketing collateral (**6%**) and sales proposals (**6%**).

This chart shows a graphical representation of the answers:



(c) Copyright 2013 DITATOO. All rights reserved.

Optional Comments

"User documentation"

"Training"

"Safety information Business Processes Internal (company-based) procedures"

"Courseware product content"

"Process documents"

"Operational standards, Guidance notes, Good practice guides"

"Engineering standards"

Restructuring and Rewriting Legacy Content (Question 5)

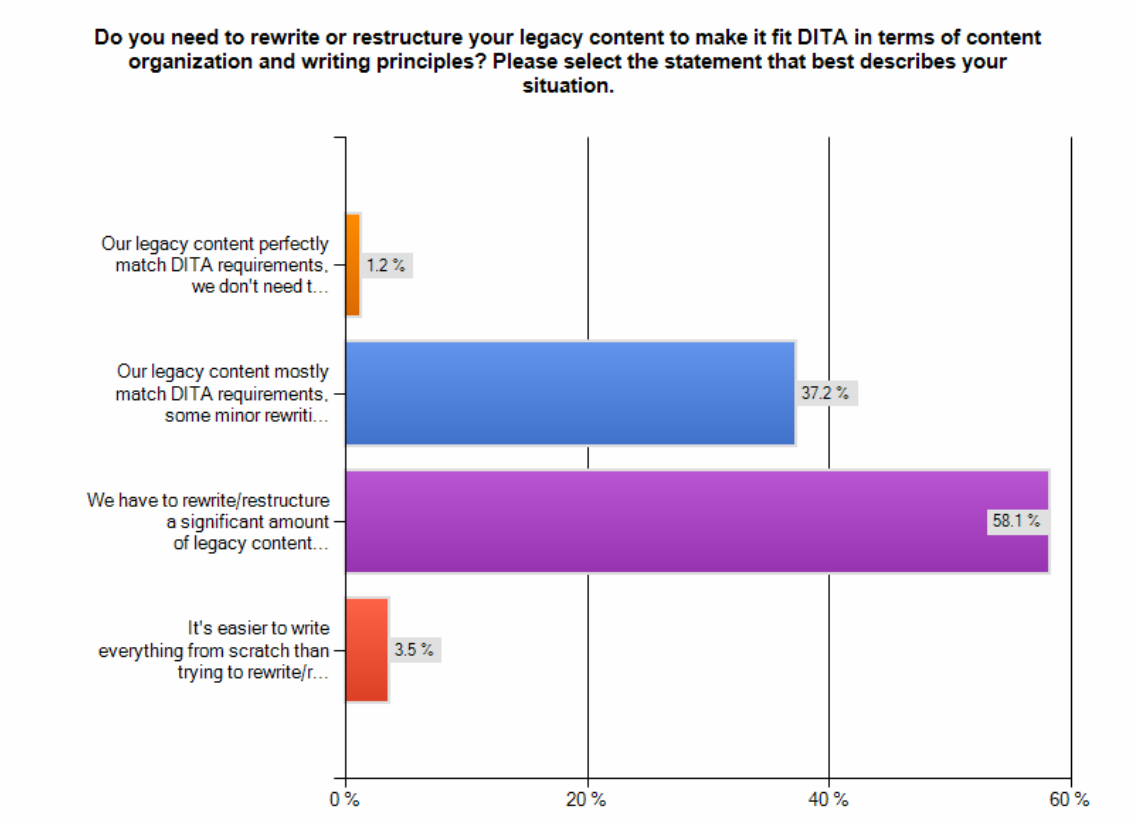
In rare cases, conversion to DITA involves just wrapping existing documents to DITA elements. Quite often, legacy documents have to be rewritten or restructured to fit DITA requirements in terms of content organization and writing principles. We asked our respondents to select the statement that best describes their situation.

The vast majority (**58.1%**) stated that they have to rewrite or restructure a significant amount of legacy content. **37.2%** answered that their legacy content mostly meets DITA requirements while some minor adjustments are still required.

Only a few respondents represented two extreme edges: **1.2%** said that their legacy documents perfectly fit DITA without any need to restructure them, and **3.5%** admitted it would be easier to write everything from scratch rather then trying to rewrite the existing content.

(c) Copyright 2013 DITATOO. All rights reserved.

This chart shows a graphical representation of the answers:


Do you need to rewrite or restructure your legacy content to make it fit DITA in terms of content organization and writing principles? Please select the statement that best describes your situation.
Our legacy content perfectly match DITA requirements, we don't need t...
1.2 %
Our legacy content mostly match DITA requirements, some minor rewriti...
37.2 %
We have to rewrite/restructure a significant amount of legacy content...
58.1 %
It's easier to write everything from scratch than trying to rewrite/r...
3.5 %
0 %
20 %
40 %
60 %

Optional Comments

"Conversion scripts took care of much of the mismatch, if more time had been spent on analysis/design of the scripts, a lot of manual cleanup could have been skipped."

"Our documents will be better formatted if we rewrite, but our import map provide loosely structured DITA."

"We aren't just re-writing for DITA; we're re-writing to improve the quality of documentated that we inherited from previous writers. Existing content has no structure, no consistency, no applied styles, and poor writing."

"We haven't absolutely decided yet. One of the goals of our pilot project(s) is to determine if we can revise/restructure before or after import. We expect to do some of both. Redundancy and inconsistency are our biggest issues - we don't really want to import everything."

"In my pilot, I choose a recent, modularised standard. (I wrote it according to some principles derived from DITA and STOP.)"

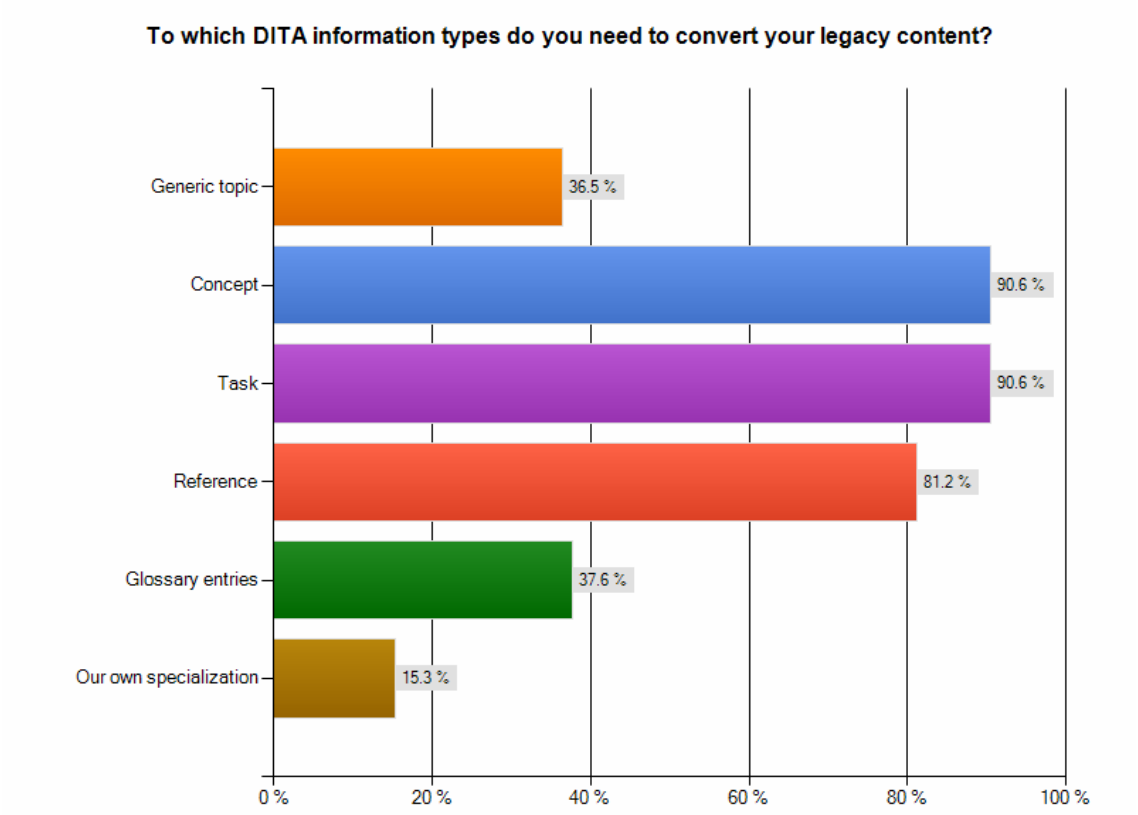
(c) Copyright 2013 DITATOO. All rights reserved.

DITA Information Types (Question 6)

We asked respondents about information types to which they convert legacy content (multiple answers could be selected).

Concept (**90.6%**), task (**90.6%**), and reference (**81.2%**) are among the leaders. Quite a lot of writers convert to glossary entries (**37.6%**) and generic topics (**36.5%**). In addition, **15.3%** convert legacy documents to a specialized information type.

This chart shows a graphical representation of the answers:



Optional Comments

"training"

"Learning&Training Specialization"

"Custom XML specification for feature sets. Generic topics may include code blocks but should be categorized separately"

"DITA L&T Specialization topics (overview, assessment, etc.)"

"Graphics"

"ditabase"

(c) Copyright 2013 DITATOO. All rights reserved.

"legal specialization will come"

"We are defining an information model."

"Can't decide if I should use the "topic" topic, or "general task". It's not computer documentation."

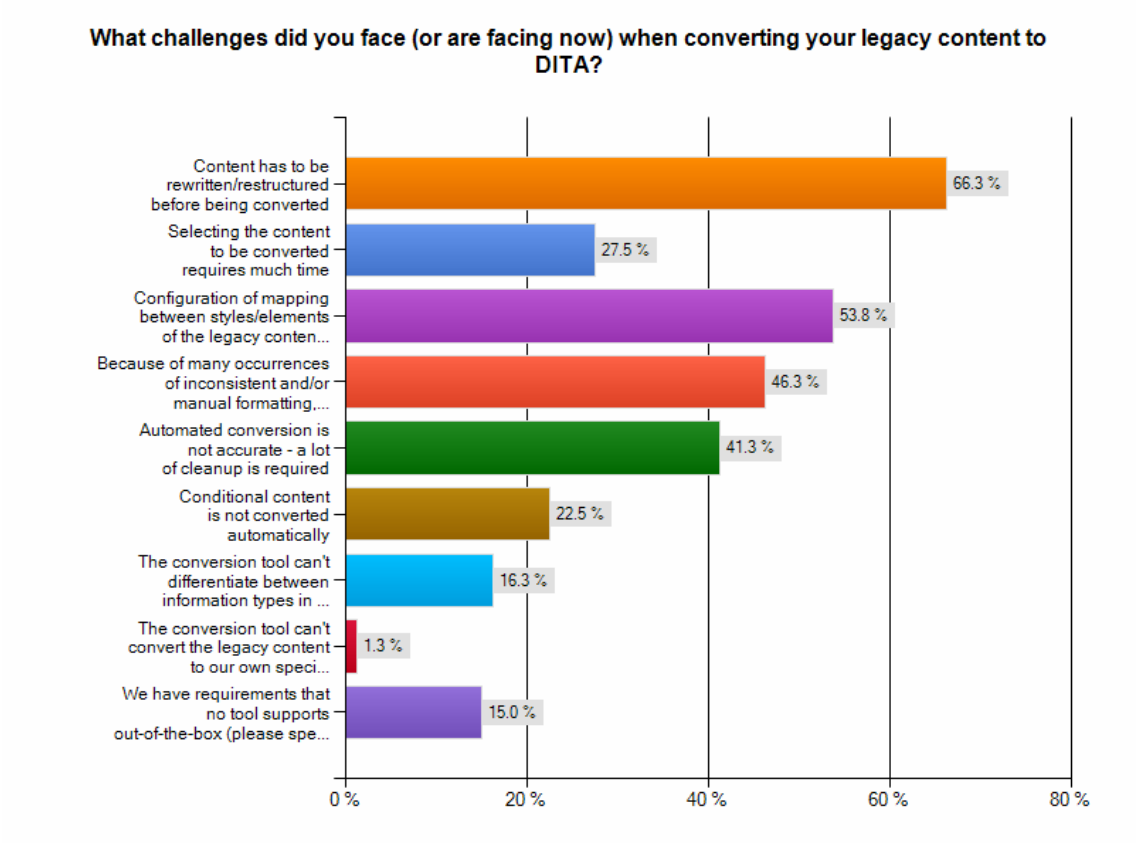
Challenges (Question 7)

To understand whether the conversion was smooth or painful, we asked our respondents about challenges they faced (multiple answers could be selected).

Considering that the majority of people answered they have to rewrite or restructure the content before the conversion (see Question 5), it comes without surprise that rewriting or restructuring the legacy content became an issue for **66.3%**.

Configuration of mapping between styles and DITA elements was an issue for **53.8%** while **46.3%** were struggling with a need to ple-cleanup legacy documents due to a significant amount of manual or inconsistent formatting.

Quite a lot of respondents (**41.3%**) stated that the automated conversion was not accurate and required a lot of cleanup. **27.5%** indicated that selecting the content to be converted requires much time. This chart shows a graphical representation of the answers:


What challenges did you face (or are facing now) when converting your legacy content to DITA?
Content has to be rewritten/restructured before being converted
66.3 %
Selecting the content to be converted requires much time
27.5 %
Configuration of mapping between styles/elements of the legacy conten...
53.8 %
Because of many occurrences of inconsistent and/or manual formatting,...
46.3 %
Automated conversion is not accurate - a lot of cleanup is required
41.3 %
Conditional content is not converted automatically
22.5 %
The conversion tool can't differentiate between information types in ...
16.3 %
The conversion tool can't convert the legacy content to our own speci...
1.3 %
We have requirements that no tool supports out-of-the-box (please spe...
15.0 %
0 %
20 %
40 %
60 %
80 %

(c) Copyright 2013 DITATOO. All rights reserved.

Optional Comments

"If there had been "real" project management, our conversion wouldn't have been such a nightmare"

"Completely unstructured content contained in PPT slides. Student Guides in Word that have little consistency from course to course."

"we have very specific requirements that no tool supported out-of-the-box: - certain pieces of content should be put into a separate file and then conref'ed - certain styles should be converted to DITA elements with the audience or product attributes set to certain values"

"Started using sophisticated conversion methods, including VI and SED, but ended up just pasting text into XMetaL tags to meet unrealistic deadlines. Managers, who had no hands-on experience at all, assumed they knew how the job should be done and could not admit that they didn't. Result was a mess."

"Less converting and more re-writing in DITA."

"We do manual conversion, works fine."

"We manually converted content from unstructured Frame to DITA. Mostly tag, then cut and paste. Hardest part was that some "topics" were a mixture of content type and has to be rewritten somewhat to separate the content into the correct topic types. Another big challange was reconditioning to fit DITA. Another challange was the graphics. We converted the WMF images (with call-outs) used in Frame to SVGZ images in DITA."

"Clarification: We rewrite/restructure after conversion, not before. Automated conversion is very good, but still needs cleanup."

"We do not use a conversion tool, but recreate the content as we decide to add it to the CMS."

"Automated conversion is not possible"

"I haven't found a conversion tool that does a good enough job with our content. We're doing it by hand which is extremely time consuming."

"These conversion problems have prevented us from pursuing automated conversion tools"

"Conversions have worked well as we have excellent developers. Some clean-up is required."

"We are expecting good success with conversion with DITATOO, but also want to use this opportunity to apply principles of minimalist authoring. Our existing documentation needs cleanup, so wholesale conversion without a target information model might be a mistake. Again, a pilot project or two will inform our decision."

(c) Copyright 2013 DITATOO. All rights reserved.

"because of the inconsistency and mixed information types, conversion scripts must be continuously updated to handle sections of the content in single publications"

"Changes are required in architectural level, and automated tools can't active it."

"Many documents have little or no recognisable structure. Legacy content has mixed information typing; modular writing is an alien concept to our SMEs. We know we will have to specialise DITA for some of our requirements."

"We leveraged existing conversion tools as much as possible (Word and FrameMaker source), but ended up writing custom code to convert our docbook source to DITA."

"Changing set element names to match paragraph names."

"Content integrity is not ensured. The current state of the proprietary xml is not necessarily representative of what it should be, therefore, a QA stage is needed."

"Having different templates, we could only map to one template. Therefore only a portion of our documents (the most current) were converted."

"We tried collaborative writing and discovered that writers are not as passionate about writing accurate topics as they are with writing their own manual."

"Cross-references are lost, which requires that we recreate them and/or set up reltables."

"Cross-references present a bootstrap problem... You need the target URLs to make them work, but we're breaking the files down from massive chapters to individual topics. Things get lost."

"DITA-OT PDF output does not have clause numbers."

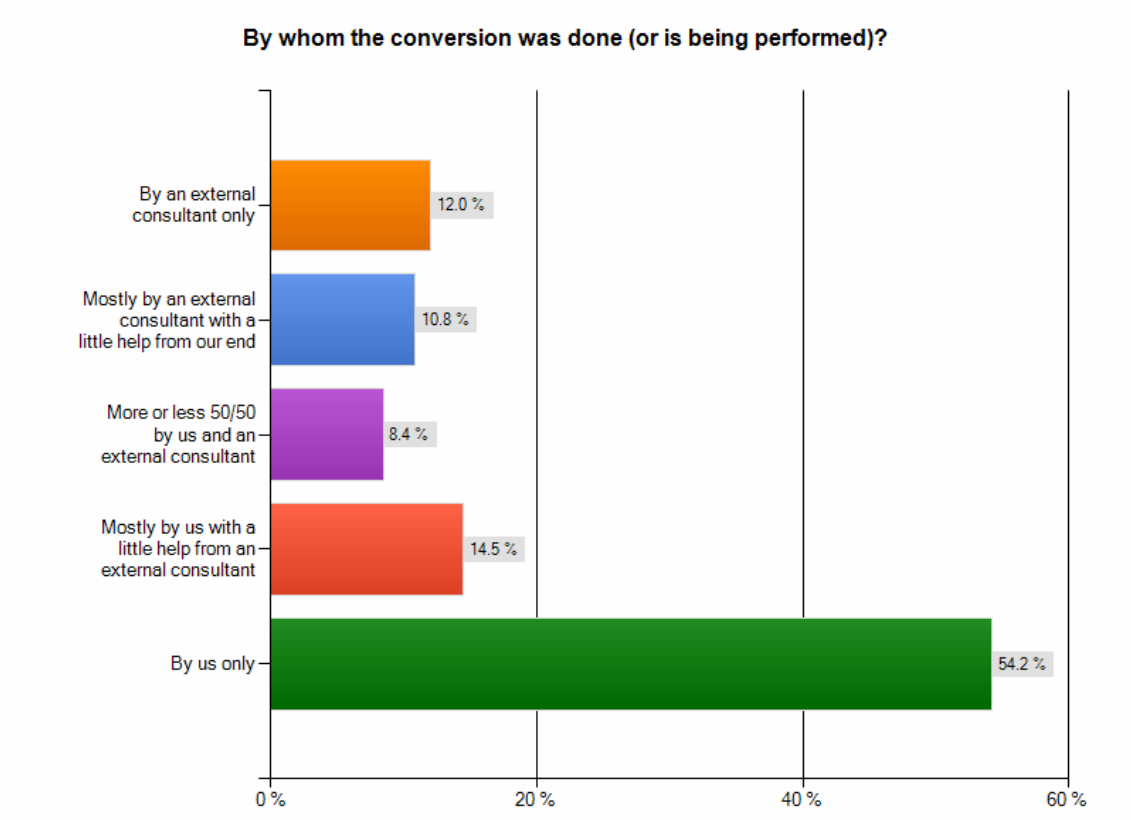
Who Was in Charge (Question 8)

We were interested to know who was in charge for the conversion. It's turned out that teams are trying to do the conversion on their own.

54.2% said that the conversion has been done completely by themselves. Some writers did most of the work by themselves with a little help of an external consultant (**14.5%**). **10.8%** answered that most of the work was done by an external consultant. **12%** indicated that the entire conversion has been done exclusively by the external consultant.

(c) Copyright 2013 DITATOO. All rights reserved.

This chart shows a graphical representation of the answers:


By whom the conversion was done (or is being performed)?
By an external consultant only
12.0 %
Mostly by an external consultant with a little help from our end
10.8 %
More or less 50/50 by us and an external consultant
8.4 %
Mostly by us with a little help from an external consultant
14.5 %
By us only
54.2 %
0 %
20 %
40 %
60 %

Optional Comments

"Replace pronouns please - "us" and "we" means me. I did all of the work."

"Smaller team included a writer transitioned to f/t developer and a writer with developer skills"

"In this business area, I am their technical writing area and I'm a consultant."

"I am their external consultant. :)"

"Not in place yet, but most likely outcome is by us with some help from external consultant."

"We hired two college interns over the summer. They completed the task in about 5 weeks."

"We pay summer students to helps us"

(c) Copyright 2013 DITATOO. All rights reserved.

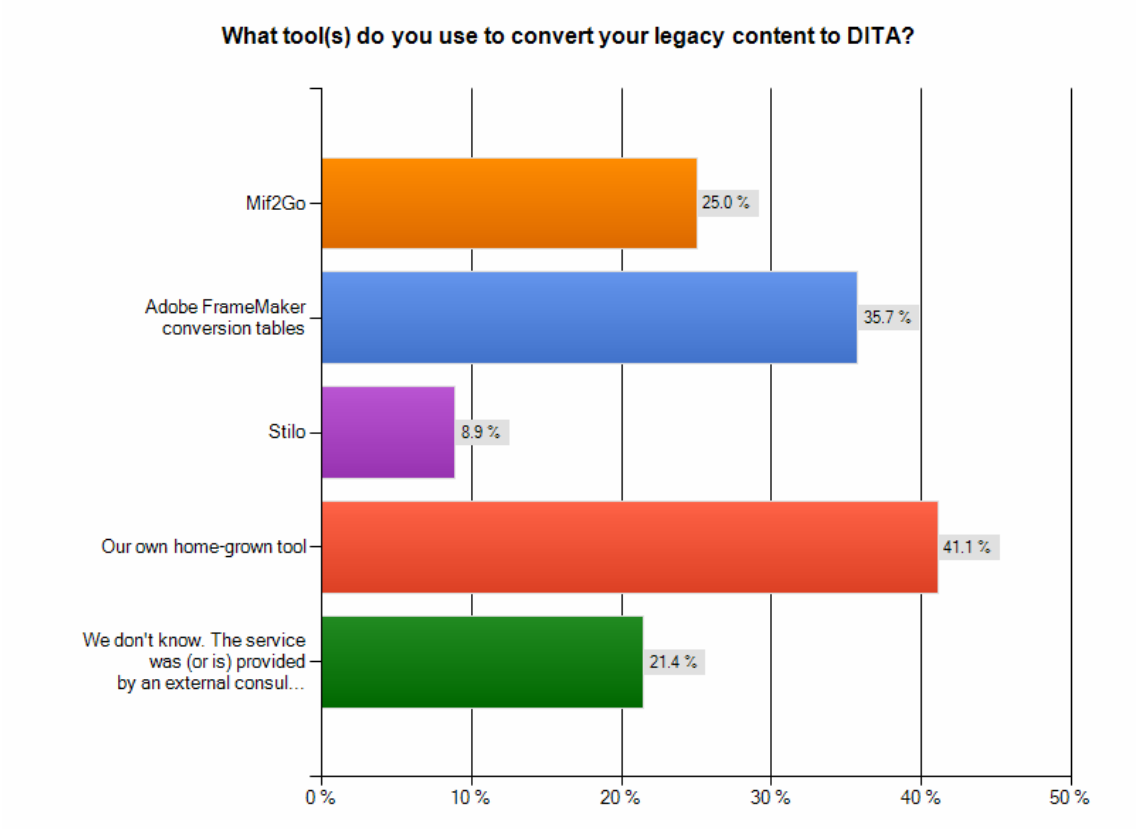
Conversion Tools (Question 9)

Conversion needs tools, and we asked our respondents about the tools they used (multiple answers could be selected).

After reading responses on the challenges our respondents faced and noticing that many comments said that very specific requirements have to be met, it comes without surprise that **41.1%** used their own home-grown tool. Quite a lot of writers (**35.7%**) used FrameMaker conversion tables and Mif2Go (**25%**). Stilo, a cloud conversion service was used by **8.9%**.

Finally, **21.4%** said they just provided their documents to an external consultant and then received the content in DITA. If we compare the latter number with the numbers received on the question about who was in charge for the conversion (Question 8), we will see that apparently, this **21.4%** is accumulated from those respondents whose external consultant was involved into the process at some level.

This chart shows a graphical representation of the answers:



(c) Copyright 2013 DITATOO. All rights reserved.

Optional Comments

"Humans"

"It's done manually - re-written"

"Manually"

"hands and brains"

"Manual copy/paste"

"Cut and pasting"

"No tool. Manual conversion."

"None. We retyped all our legacy content."

"Copy and paste plus rewriting a lot of content"

"We did it manually since out product changed greatly during the conversion."

"We did not use any tools. Instead, we used the conversion exercise as the adoption path for writers."

"We will not convert, rather we will restructure and rewrite the content."

"We copy and paste manually from Word to Oxygen."

"oXygen"

"I'm rewriting, restructuring and converting to DITA via oXygen 14"

"DITAToo"

"We will be using DITAToo for conversion and content management, and authoring and revising in structured FM."

"None. We're doing it with the DITA authoring tool embedded with our CCMS system."

"Home grown method: Create a map-and-topic structure to match the table of contents of the original -- but with empty topics. Then copy-and-paste-and-restructure-and-rewrite-on-the-fly. This seems to be the process most people use."

"Arbortext Import"

"Arbortext Editor"

"Conversion was done by me using XSLT 2.0/saxon 9.x"

(c) Copyright 2013 DITAToo. All rights reserved.

"We hired Rick Quatro to do the conversion tables and he wrote FrameScripts for us to automate the conversion to DITA XML. He did the initial conversions, but now we do them in house using his scripts."

"The external consultant carried out the migration to new DITA environment, and we made the change from previous info types to DITA info types."

"We created our own conversion tools using Ant and XSL. For unstructured FrameMaker, we first converted to a structured version of the unstructured template and exported that to XML."

"Content Mapper has a useful import feature"

"You need to include Document Control Laboratories, a major supplier of conversion services"

"Conversion tables in Maker, then Maker scripting or FDK programming to break out individual topics and fix the xrefs."

"I'm hoping the conversion tables do the trick for us. Yet to be done."

"Flattened DocBook books. docbook2dita.xsl transformation from the DITA Open Toolkit plugin, run in oXygen. General cleanup."

"Export text from PDF with Adobe Acrober text exporter (it doesn't retain line breaks.) Emacs for clean up. Bascailly rewrite in DITA using this raw text. Not as hard as it sounds, because DITA is quite simple to use."

"Word to Flare, Flare to DITA"

"Madcap flare"

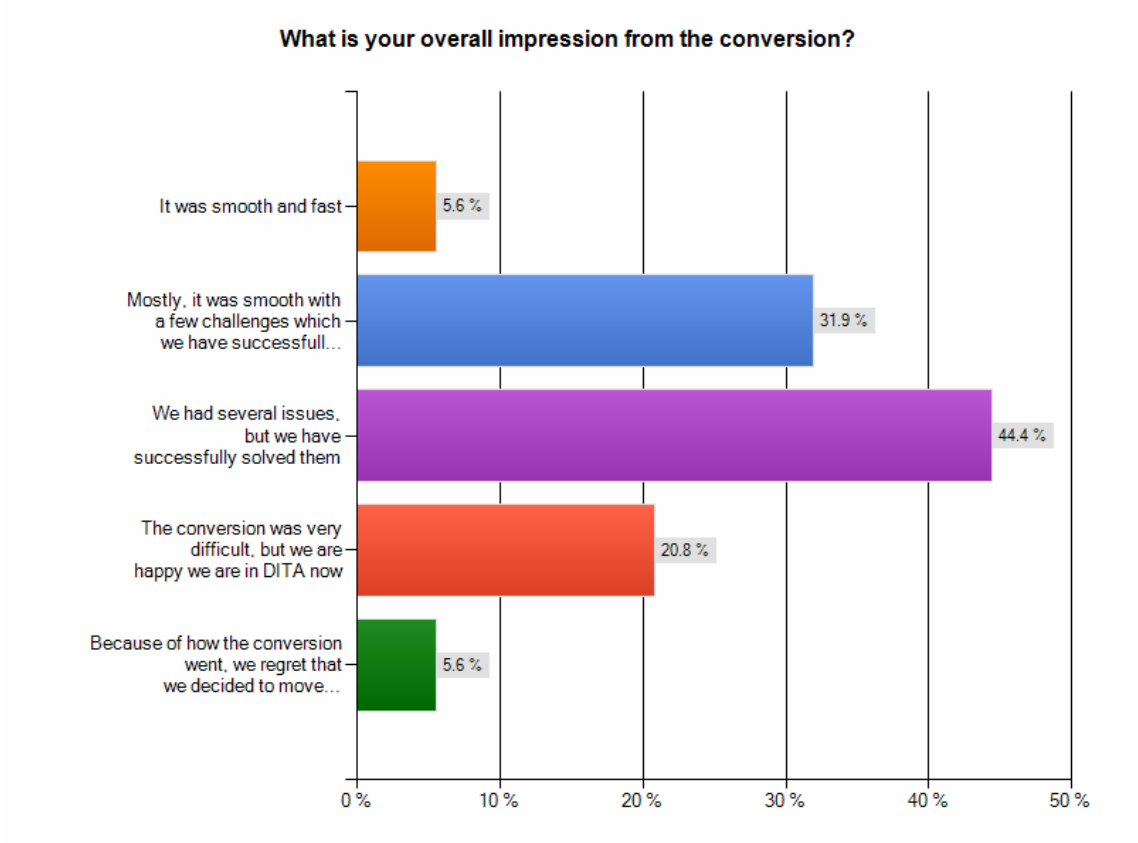
Overall Impression (Question 10)

We asked about an overall impression that our respondents gained from the conversion (multiple answers could be selected).

Most of them (**44.4%**) answered that they did have several issues, but these issues have been successfully resolved. **31.9%** indicated that the conversion went smoothly with a few challenges. **20.8%** said the conversion was very difficult, but now, when they are already in DITA, they are happy. Finally, **5.6%** admitted that because of how the conversion went, they regret they decided to move to DITA. The same number exactly answered that the conversion was smooth and fast.

(c) Copyright 2013 DITATOO. All rights reserved.

This chart shows a graphical representation of the answers:


What is your overall impression from the conversion?
It was smooth and fast
5.6 %
Mostly, it was smooth with a few challenges which we have successfull...
31.9 %
We had several issues, but we have successfully solved them
44.4 %
The conversion was very difficult, but we are happy we are in DITA now
20.8 %
Because of how the conversion went, we regret that we decided to move...
5.6 %
0 %
10 %
20 %
30 %
40 %
50 %

Optional Comments

"I converted a 250 complex technical manual as a pilot. I throught it was straightforward, but senior management did not understand the workflow at all."

"The company direction is DITA, they wasted a huge amount of money by not going with Stilo and by not spending the time to figure out what was really required."

"I would never have done it, despite the fact that I know a lot about XML/DITA. It was not the correct call for that business unit, but my opinion was not important."

"The documentation is in very bad shape. Automated conversion will not work correctly."

"It is too early to tell, but it is going to be a long slog."

"Legacy conversion has not been easy in any case. Some groups wonder why we are doing this. Many others find a lot of value."

"Doing manual allows to fix issues just at the right time"

"Mostly, it was smooth with a few challenges, that we will attempt to solve for the next

(c) Copyright 2013 DITATOO. All rights reserved.

document."

"PDF output - bought RenderX to minimize"

"Since we created our own tool, we could refine the conversion logic for each set of content until we were satisfied. This reduced the amount of postprocessing necessary."

"Well, we're optimistic that we'll solve the issues."

"We aren't at this point yet, but I expect to be satisfied AND have some work to do! We aren't looking for quick and easy - we need to build a better product and serve a wider range of information consumers than we currently do."

"Still in the planning stages, but hope to end up with option 4. It will be tough, but worthwhile."

What If... (Question 11)

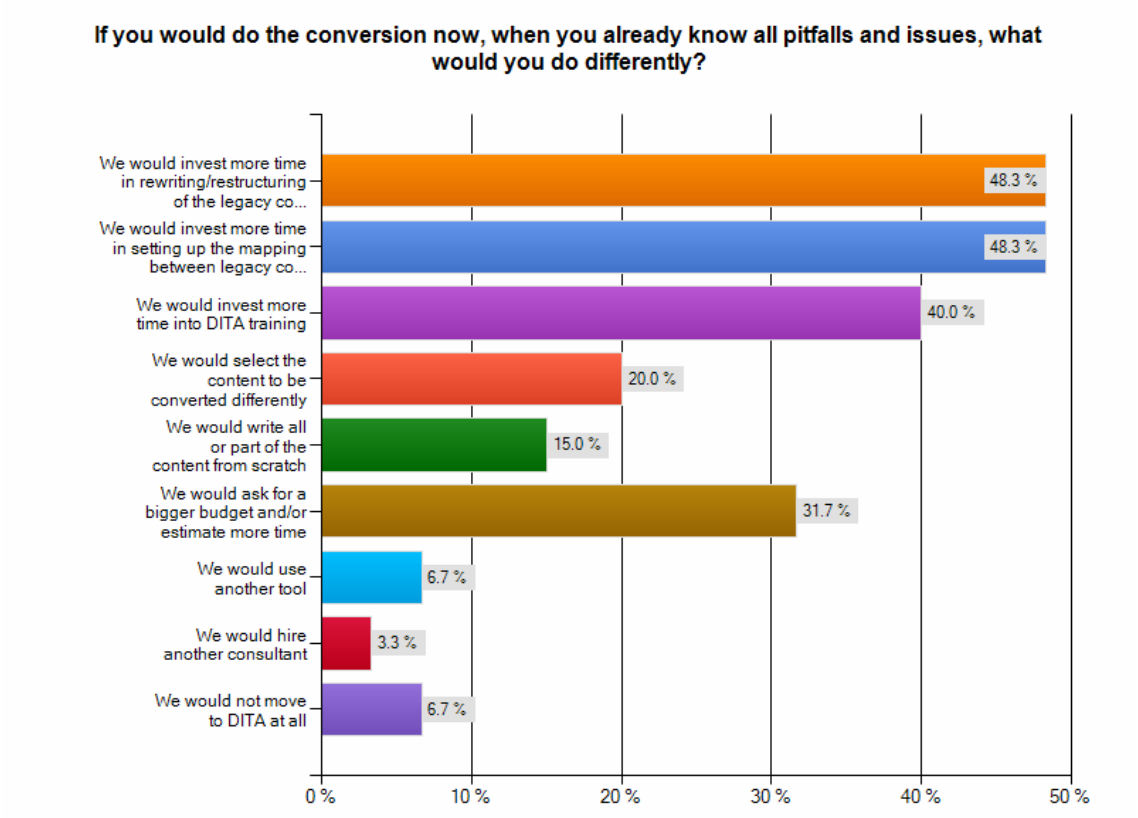
There is always "what if" in such cases. So we asked our respondents if they would do the conversion now, when they already know all pitfalls and issues, what they would do differently (multiple answers could be selected).

Here are the leaders:

1. "We would invest more time in rewriting/restructuring of the legacy content": **48.3%** (compare it with **66.3%** of people who named the need to restructure or rewrite the legacy content as a major challenge in Question 7)
2. "We would invest more time in setting up the mapping between legacy content's styles/elements and DITA": **48.3%** (compare it with **53.8%** of people who named the mapping as a major challenge in Question 7)
3. "We would invest more time into DITA training": **40%**
4. "We would ask for a bigger budget and/or estimate more time": **31.7%**
5. "We would select the content to be converted differently": **20%** (compare it with **27.5%** of people who named selecting the content to be converted as a major challenge in Question 7)

(c) Copyright 2013 DITATOO. All rights reserved.

This chart shows a graphical representation of the answers:


If you would do the conversion now, when you already know all pitfalls and issues, what would you do differently?
We would invest more time in rewriting/restructuring of the legacy co...
48.3 %
We would invest more time in setting up the mapping between legacy co...
48.3 %
We would invest more time into DITA training
40.0 %
We would select the content to be converted differently
20.0 %
We would write all or part of the content from scratch
15.0 %
We would ask for a bigger budget and/or estimate more time
31.7 %
We would use another tool
6.7 %
We would hire another consultant
3.3 %
We would not move to DITA at all
6.7 %
0 %
10 %
20 %
30 %
40 %
50 %

Optional Comments

"We would have begun in DITA in the first place."

"I believe that moving to DITA is sometimes a correct decision, but it has caused an irrational stampede in the techcomm community. Writers and managers use it as an opportunity to complicate their jobs so that they can appear to be indispensible."

"We hired a consultant. Aldousxml.com"

"Had this project happened now, we would have tried DITA4Publishers to export from Word to DITA"

"I would get a better understanding of XSLT to create nice PDF output."

"We would ask for much more training"

"No, it also train people"

"Will probably regret not estimating more time, investing in DITA training, setting up mapping etc."

(c) Copyright 2013 DITATOO. All rights reserved.

"Allocate resources and time for analyzing the content and developing the target information architecture and reuse model. Then develop the conversion process to match that target model."

"Invest more time on data modelling/analysis"

"We are constantly converting Word docs from Engineering so it's an ongoing task. We have it down to a science now, so we wouldn't change unless were made aware of something more powerful than MiF2Go."

"costs are regularly underestimated because the problems grow as you get into the details of the content"

"More test runs to catch all the gotcha's before the global conversion."

"Again we are just beginning our project. We spent months in product selection and are now at the educating phase. We've been very careful in our planning and I hope that in the end I would not check any of the above!"

"Do not know yet, but can be almost certain that a bigger budget and more time will be needed."

"Would introduce more style consistency in the source. Would elaborate the embedded processing instructions to get a better DITA output. Otherwise it's pretty easy."

"I think we go the same basic route."

"I knew what to expect before starting. Management of the subject matter expects is the real issue for us, plus a huge volume of legacy material and an uninterested board."

"None. We were aware of most of the pitfalls going into the conversion."

"Nothing"

" No significant changes"

(c) Copyright 2013 DITATOO. All rights reserved.

Final Comments (Question 12)

Finally, we asked our respondents to share with us any comments or thoughts that they have on the conversion. These are the comments that we have received:

Optional Comments

"I like the DITA format, but it's never going to be adopted in my organisation. It's just not flashy and shiny enough for them, and they have have no experience of anything but Word. It's a shame because we have lots of overlapping content that is hard to manage, and DITA's topic/reltables/conref/conditional profiling ideas are a really good fit for us."

"Now that I've been through a RoboHelp to Arbortext-based, IBM DITA conversion, I'd love to be a consultant and help other IBM acquisitions get through this more easily!"

"Move to DITA sometimes has to be done to reduce localization costs, but often that can be achieved by other means. In my case, it could have been done by using Madcap Lingo (and a little advance planning). However, my company had a separate localization unit that was worshipped by upper management, who poured so much money into it that they were locked into an unwise decision made years before. Writers were not empowered to participate in these decisions, much less the final appearance of their single source publications. Years before doing the conversion, I had a feeling that the integrity of the content could be damaged by moving to DITA, but kept an open mind and just went ahead and did it. But in hindsight, I was right. DITA documentation takes such a long time to do that the entire nature of a technical writer's job is changed. Mechanical processes and formats that are determined by others are the main concerns, and content is taken completely out of the equation. After DITA I no longer had time to do the research that the apps required, and both documentation and products suffered because of that. But I did resolve the problem - by leaving the company. :-)"

"I regret that my experience with DITA has only been in the conversion end, and have not had the experience working with DITA workflows in progress, so I don't have any basis for comparison."

"If new to DITA, I feel it is best to perform the conversion on my own. This helps me test my new learning, I feel like I own the content in DITA and it is cost effective too. Of course, we might be compromising on time a bit but it is better than spending time on correcting stuff that was converted by a tool or an external agency."

"The quality of the source is the most important factor in the success of the conversion/migration effort. We have created automated end to end migration tools which combine content export, conversion, validation, import etc which help enormously."

"Conversion" for us has two possible meanings/stages. 1- conversion of systems 2- conversion of content. We created a new content ecosystem to replace an older proprietary one. We were given 3 months to do replace end to end (authoring to pdf/epub publishing).

(c) Copyright 2013 DITATOO. All rights reserved.

Our initial content creation did not involve conversion of legacy content, but rather the development of new content for products. This was a benefit, because it gave me a better (not perfect) target to map to when developing the transform from legacy xml to dita xml. It would have been much more difficult if we had to convert legacy content to DITA first, without properly understanding the target models. If we had to take that approach, I'm positive that many mistaken assumptions regarding the model would have been made, requiring constant rewrite/re-run of the transforms."

"We are writing a lot of new technical information from scratch directly in DITA but our user manuals are still in Word and we plan to convert them by copying and pasting unless we find a tool to do it automatically."

"We were already topic-based and very structured with our Frame docs (but not using Structured FM). This made the mapping to DITA XML elements very straight forward. There is some cleanup, but most of the post-conversion work is setting up the newly converting files into our CMS tool."

"We are not sure yet about the best way to handle a) our mixture of referenced and embedded graphics b) screenshots with callouts."

"More thorough analysis before starting the product would habe been fine. It does not help to have your content converted but still no standardized manual as to chapter structure and content topics."

"The conversion is actually going very smoothly. I know the content and can identify where to reuse chunks. The client is happy (so far) with the approach (single source vs copy and paste approach to content management) and the look and feel of the output, not to mention the multiple (filtered) versions of the document."

"We converted legacy content as needed. Needed meaning that we want the content in DITA for future editing or translations. Legacy content that rarely changes or does not get translated will remain in Frame for now. The lagacy content that we did convert to DITA was poorly structured for the most part and we used the DITA framework to improve the quality and consistancy of the content. Overall our content has significantly improved and we have gone from problematically publishing in Frame (every publication has some error or two) to publishing 100% error free in DITA. And with the CMS, our editorial time for DTIA has been greatly improved in that we can now verify that all content has been approved at publishing time. Moving to DITA was the right choice for us. We currently have about 40 bookmaps that produces about 150 different publications. An that is just for our product user manual. We also have many product e-learning modules that are published in both ASP and PDF formats using the same source files."

"The determining factor is often: "Is our legacy content already topic-oriented?" If your legacy content tends to contain long topics containing several sections and subsections, mixing reference and task content together, you're going to have a lot of rewriting to do before you convert. If, however, your legacy content is already designed to be granular,

(c) Copyright 2013 DITATOO. All rights reserved.

you'll have a much easier time. If your legacy content is not topic-oriented, but will not evolve much (or at all) after conversion, consider converting it to the generic topic type instead of forcing yourself to rewrite it."

"We have not met a Word use who perfectly stylese documents. We have found success even with some inconsistent styles."

"I wrote a lot of custom tools to do the job. They don't work great but they work better than the off the shelf packages I saw."

"We are just getting started, so my answer to question 11 is not really accurate. It is just my impression so far."

"Converting to DITA requires time and resources. Difficult when there are ongoing scrum meetings, and releases. Have to block time off and require discussion and understanding from development that documentation issues are not handled immediately. Have to draw the line what gets converted or not and for which product lines. Again time and resources!"

"I'm excited at this point!"

"We have not yet started, but are in the planning stages. Our workflow is going to involve people writing in Word for some time, so there will always be conversion. The challenges will be in getting people to write in a modular way, and to adhere to limited quasi-structured templates that make the conversion process as easy as possible."

"I like the DITA format, but it's never going to be adopted in my organisation. It's just not flashy and shiny enough for them, and they have have no experience of anything but Word. It's a shame because we have lots of overlapping content that is hard to manage, and DITA's topic/reltables/conref/conditional profiling ideas are a really good fit for us."

"The process has gotten easier over time. The biggest advances were Mif2go and dita2go. Very few legacy projects are written in the strict topic formats required for DITA and that is biggest challenge. I do not see how a tool can deal with that aspect of a conversion from Word or even FrameMaker."

"Thanks for the survey."

About DITATOO

DITATOO is an affordable lightweight DITA Content Management System (DITA CMS). While DITATOO has a clean and straightforward user interface, it still provides a powerful solution for automating and facilitating many of the DITA-related tasks that you have to do everyday.

This includes metadata-based and full-text search, link management, capabilities to arrange topics into deliverables directly in DITATOO, publishing to a wide variety of formats, managing conditional publishing, translation management, version control, and much more.

For more information about DITATOO DITA CMS, go to http://ditatoo.com/.

(c) Copyright 2013 DITATOO. All rights reserved.